DOCUMENT RESUME

ED 229 432 TM 830 334

AUTHOR Gastright, Joseph F.
TITLE An Experimental Study of the Effects of Intensive Test Wiseness Instruction on the Reading Gains of Disadvantaged Elementary School Pupils.
PUB DATE Apr 83
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Compensatory Education; *Educationally Disadvantaged; Elementary Education; Intentional Learning; Pretests Posttests; *Reading Improvement; *Reading Instruction; Reading Tests; Research Methodology; Surveys; Teacher Attitudes; *Test Coaching; *Test Wiseness
IDENTIFIERS California Achievement Tests; Elementary Secondary Education Act Title I; Mini Tests

ABSTRACT

This study tested the hypthesis that the reading gains of compensatory education pupils in grades one and four exposed to intensive test wiseness instruction would not differ significantly from the reading gains in a randomly assigned control group. Schools offering ESEA Title I classes (N=48) were randomly assigned to the treatment and control groups. Approximately four hundred students in each grade received twenty hours of test wiseness instruction as part of one hundred hours of supplemental reading instruction between the pretest and the posttest administration of the California Achievement Test. The implementation procedures of RMC Model A were rigorously observed and the NCE gains of the school units were analyzed separately for the two grades using a t test. The null hypothesis was supported in both cases. Extensive evidence was collected to show that the materials were used by the treatment teachers. Survey results indicated that teachers were quite positive about the commercially prepared mini-tests and felt that they would improve reading gains. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED229432

An Experimental Study of the Effects of Intensive Test Wiseness Instruction on the Reading Gains of Disadvantaged Elementary School Pupils

JOSEPH F. GASTRIGHT
Cincinnati Public School System

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. F. Gastright

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the annual meeting of the
American Educational Research Association
Montreal, Canada
April, 1983

TM 830 334

An Experimental Study of the Effects of Intensive Test Wiseness Instruction on the Reading Gains of Disadvantaged Elementary School Pupils

JOSEPH F. GASTRIGHT
Cincinnati Public Schools

Test wiseness as a measurement concept has grown from a proposed inhibitor of reliability (Thorndike, 1951) to a delineated taxonomy of test taking principles (Millman, et.al., 1965). As test wiseness has gained acceptance as a mediator of the accuracy of test scores, researchers have reported successful programs that improve test wiseness (Gains and Jongsma, 1974; Oakland, 1972; Slakter, et.al., 1970; Callenbach, 1973). Recently commercial houses have offered instructional materials for use in elementary and secondary schools. Shuller (1979) reported that significant gains in reading were made by pupils in schools using published test wiseness programs.

The empirical evidence supporting test wiseness instruction is modest and occasional studies have reported "non-significant" differences (Jongsma and Warshauer, 1975). No solid evidence has been presented on effects of using test wiseness with the disadvantaged students in compensatory education programs. Diamond, et.al. (1976) provided non-experimental results which suggested that inner city pupils in grades five and six may have better test wiseness skills than they are credited with by their teachers. The authors speculated that test wiseness instruction in a practical setting would probably not produce dramatic improvements in scores. The objective of the present investigation was to assess the effects of intensive test wiseness instruction on the standardized reading test gains of first and

fourth grade students in a compensatory education program. The study tested the hypothesis that the reading gains of pupils exposed to intensive test wiseness instruction would not differ significantly from the gains in a randomly assigned control group.

Methods

The study was conducted by a large urban school district which offered compensatory reading programs for first through fourth graders in public and non-public schools. Forty-eight schools offering these programs were randomly assigned to one of two groups. The first group used the Mini-tests (Hinman, et.al., 1970) in grade one while the fourth grade students in those schools served as a control group for the second group. The second group of schools used the Mini-tests in grade four while the first grade students served as a control group for the first group.

All students in the experimental and control schools were pretested within one week of the empirical norm date in fall with the total reading section of the California Achievement Test. The students were posttested in spring within one week of the empirical norm date. Level 11, form C was used in the first grade and Level 14, form C was used in the fourth grade. It was impossible in the field setting to guarantee the number of students selected for the program in advance. It was arbitrarily determined that schools with fewer than five pupils at a grade would be eliminated from the study. All scores were converted to normal curve equivalents and mean school gain scores for each grade were analyzed using a t test.

Fourth grade students were eligible for the compensatory education class if they scored at or below the 41st NCE on the total reading section of the California Achievement Test administered during the previous April. First

grade students who scored at or below the 33rd percentile on the Metropolitan Readiness Test (1976 Edition) Level II, form P were eligible.

Compensatory education students in both the treatment and the control group received 50 minutes a day additional reading instruction in small groups of six to eight. Teachers in the treatment group used approximately 40 Mini-test exercises following the teacher instructions provided with the materials. The teachers were introduced to the materials in a workshop and their use of the materials was monitored by project supervisors.

The test wiseness materials emphasized accuracy, speed and endurance in test taking. The exercise sheets featured a variety of item types in vocabulary and reading comprehension. The materials concentrated on many commonly mentioned test wiseness skills such as paying attention to directions, reading all answer choices, eliminating wrong answers and pacing. In contrast to other commercial materials, the Mini-tests overtly instructed the teacher to help students identify "traps in the test" and to try to develop insights into the characteristics of commonly used test item distractors. Individual exercises utilized item types and formats identical to those contained in the California Achievement Test. The exercises were used throughout the six months between testings and the teachers were encouraged to provide review activities during the period immediately before the administration of the posttest.

Any school with gain scores on fewer than five pupils was excluded from the final analysis. In the first grade the test wiseness materials were used by 61 compensatory education teachers with 321 students in 19 schools. Four schools were eliminated from the first grade sample because of low numbers. The first grade control classes included 395 students in 21 schools. Two schools were excluded from the first grade control group because of low numbers.

In the fourth grade the test wiseness materials were used by 64 compensatory education teachers with 405 students in 20 schools. Three schools were eliminated from the fourth grade sample because of low numbers. The fourth grade control classes included 209 children in 11 schools. Four schools were eliminated from the fourth grade control group due to low numbers.

A further 9 control schools disappeared from the random sample at the fourth grade for a reason not anticipated by the author. Schools were free to select the grades served by the compensatory education teachers at the local level. Schools which were notified that they would receive test wiseness materials in the fourth grade organized their classes to include fourth grade students. Schools which were notified that they would receive first grade materials chose to use their staff in the lower grades and in 9 cases, did not operate fourth grade classes. The problem did not occur in first grade because all schools were strongly urged to include all eligible first graders for obvious educational reasons.

Results

The reading achievement gains and standard deviations of the treatment and control groups are given in Table 1. The difference between school unit gains was 3.1 NCE's at the first grade (Treatment 9.7; Control 6.6). The

(INSERT TABLE 1)

difference in first grade means was not significant ($t = 1.06$, $df = 1, 38$). The difference between school unit gains was 2.0 NCE's at the fourth grade (Treatment 7.1; Control 5.1). The difference in fourth grade means was not significant ($t = 1.00$, $df = 1, 29$). The null hypothesis was supported in both cases.

Inspection of the variance in the first grade suggested that the assumption of homogeneity variance might have been violated. Despite the often mentioned robustness of the t statistics, the F_{max} procedure suggested by Winer (1962) was carried out and the hypothesized homogeneity of variance was rejected ($p < .01$). The null hypothesis at grade one was retested using the t' statistics developed by Welch and discussed by Winer (1962) as perhaps the soundest technique for testing hypotheses about the difference of two means when the sample variances are found to be non-homogeneous. The null hypotheses for grade one was again supported ($t' = 1.03$, $df = 1, 27$).

(INSERT TABLE 2)

Discussion

The findings of this field experiment do not support the prevailing notion that intensive test wiseness instruction will produce higher test scores for poor readers in economically disadvantaged schools. The pretest scores were well below average. The mean pretest score for the first grade group was at the 16th percentile on the test norms and the mean pretest score for the fourth grade group was at the 15th percentile.

In this study, inner city students in both the treatment and control schools made norm-referenced gains which ranged from .24 to .46 of an inter-student standard deviation on the outcome measure. Gains in this range are not unusual with Fall-Spring Testing in compensatory education programs. The mean differences between the treatment and control groups averaged 2.5 NCE's or about .12 of an inter-student standard deviation. Differences on this order would be well within the reported range of gains attributed to "coaching" on various standardized tests.

While test wiseness skills were originally proposed as a set of content free strategies for optimizing test performance, the Mini-tests used in this study contain a sizeable dose of reading content and vocabulary which could be used quite effectively in reading instruction. It is possible that the positive achievement gains noted in other studies using similar materials were due as much to the added reading instruction as to the development of independent test wiseness skills. Certainly the materials provide exposure to the more difficult and discriminating types of thinking required in reading comprehension. If additional content exposure explains reported test wiseness gains then it is possible that the gains would be minimized in the present population of students, all of which received extensive additional reading instruction in small groups.

There is the additional possibility that the double exposure of compensatory education students to annual pretests and various diagnostic tests has increased their test wiseness to the point where deficiencies in this area no longer significantly lower observed test scores. An important question not raised by test wiseness investigators is whether above average test wiseness skills actually mask content area weaknesses on norm referenced achievement tests. To the extent that a subject is able to score higher on a test due solely to content free skills, then the information contained in his/her normative test scores does not accurately reflect achievement in the content area. Is it possible to have too much test wiseness knowledge on a norm referenced test?

In addition to the regular classroom reading instruction, pupils in both the treatment and control group received approximately 100 hours of supplementary reading instruction in ESEA Title I classes between the pre-

test and posttest. In the treatment group approximately 20 hours of this instruction was based on the Mini-test series.

Additional data was collected to assure that the treatment was being administered. Teachers using the test wiseness materials were required to keep a log of the date the materials were used and to rate each of the Mini-tests on their utility as instructional materials. Teachers were encouraged to use the materials as part of their regular supplementary instruction and to follow the publisher's directions closely. Title I supervisors observed the use of the materials in the classroom and reported that the teachers' reports of usage were accurate. The teacher logs indicated that virtually all of the treatment students received instruction on 75 percent or more of the exercises.

During mid-April, a teacher survey was sent to each teacher who had participated in the use of the test wiseness materials. Of the 134 teacher participants, 111 or 83 percent completed and returned the survey. With the exception of the difficulty level of the first grade materials, all survey items were rated quite positively by the teaching staff. In general, fourth grade teachers were more positive than teachers in the first grade.

Hundreds of written comments regarding various uses, problems and strengths of the test wiseness materials have been summarized into a few general categories. The materials were credited with being very useful in developing dictionary and listening skills as well as improving attitudes toward testing and following directions. Many teachers felt that they could pinpoint specific problems, diagnose and remediate them better as a result of the test wiseness material. Other teachers liked the format

and high interest level of the materials. The ease of using prepared ditto sheets and the fact that all of the materials were available for use was also mentioned repeatedly as a strength of the program. There was a pervasive belief that test taking would be improved as a result of the test wiseness efforts.

Despite the positive response of teachers, this study found no significant improvement in the reading gains of students instructed with the test wiseness materials. The results are in accord with Diamond's prediction (1976) that scores would probably not rise dramatically following test wiseness instruction in practical testing situations. The continued support of test wiseness instruction by the treatment teachers and the widespread belief that inner city children need this kind of help, supports a contention that the affective response of pupils and teachers to test taking may be the real issue in the use of test wiseness materials.

Additional controlled studies using a variety of test wiseness materials and outcome measures in various field settings are needed before final determination can be made on the viability of the construct. This study adds one piece of information to the debit side of the scale.

Table 1. Reading Achievement Gains, Standard Deviations and Sample N's for the Test-Wiseness and Control Group at Grades One and Four.

Grade	Test Wiseness	Control
One		
$\bar{x}$	9.67	6.61
S.D.	11.70	5.95
N	(19)	(21)
Four		
$\bar{x}$	7.14	5.10
S.D.	5.48	5.36
N	(20)	(11)

Table 2. Summary Table on the Application of Welch's (1947) t' Test for the Hypothesis That $\mu_T - \mu_C = 0$

Statistic	
$t' = \dfrac{(\bar{x}_T - \bar{x}_C) - (\mu_T - \mu_C)}{\sqrt{(s_T^2/N_T) + (s_C^2/N_C)}}$	= 1.03
$C = \dfrac{s_T^2/N_T}{s_T^2/N_T + s_C^2/N_C}$	= 0.81
$f = \dfrac{f_T f_C}{f_C C^2 + f_T (1-C)^2}$	= 27.14

References

Ayrer, J., Diamond, J. J., Fishman, R., & Green, P. "Are inner city children test-wise?" Journal of Educational Measurement, 1976, 14, 39-45.

Callenbach, C. The effects of instruction and practice in content-independent test-taking techniques upon the standardized reading test scores of selected second grade students. Journal of Educational Measurement, 1973, 10, 25-30.

Gaines, W. G., & Jongsma, E. A. The Effect of Training in Test-Taking Skills on the Achievement Scores of Fifth Grade Pupils. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois, April, 1974.

Hinman, D. E., Mitchell, K. A., Shuller, S. M., & Schwartz, T. F. Mini-Tests. New York: Educational Solutions, Inc., 1976.

Jongsma, Eugene A., and Elaine Warshauer. The Effects of Instruction in Test-Taking Skills upon Student Performance on Standardized Achievement Tests. Final Report. Baton Rouge, La.: Louisiana State University, 1975.

Millman, J., Bishop, C. H., & Ebel, R. An analysis of test-wiseness. Educational and Psychological Measurement, 1965, 25, 707-726.

Oakland, T. The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. Journal of School Psychology, 1972, 10, 355-360.

Slakter, M. J. The effect of guessing strategy on objective test scores. Journal of Educational Measurement, 1968b, 5, 217-221.

Shuller, Stephen M. A Large-Scale Assessment of an Instructional Program to Improve Test-Wiseness in Elementary School Students. New York, N.Y.: Educational Solutions, Inc. 80 Fifth Ave., New York, NY 10011, 1979.

Thorndike, E. L. Reliability. In E. F. Lindquist (Ed.), Educational Measurement. Washington, D. C.: American Council on Education, 1951.

Winer, B. J. Statistical Principle's in Experimental Design. New York: McGraw-Hill, 1962

ABSTRACT

An Experimental Study of the Effects of Intensive Test Wiseness Instruction on the Reading Gains of Disadvantaged Elementary School Pupils

This study tested the hypothesis that the reading gains of compensatory education pupils in grades one and four exposed to intensive test wiseness instruction would not differ significantly from the reading gains in a randomly assigned control group.

Schools offering ESEA Title I classes (N = 48) were randomly assigned to the treatment and control groups. Approximately four hundred students in each grade received twenty hours of test wiseness instruction as part of one hundred hours of supplemental reading instruction between the pretest and the posttest administration of the California Achievement Test. The implementation procedures of RMC Model A were rigorously observed and the NCE gains of the school units were analyzed separately for the two grades using a t test. The null hypothesis was supported in both cases. Extensive evidence was collected to show that the materials were used by the treatment teachers. Survey results indicated that teachers were quite positive about the commercially prepared mini-tests (Hinman, et al; 1970) and felt that they would improve reading gains.